

COVID-19 Register; Quality Statement

Exported from METEOR (AIHW's Metadata Online Registry)

© Australian Institute of Health and Welfare 2024

This product, excluding the AIHW logo, Commonwealth Coat of Arms and any material owned by a third party or protected by a trademark, has been released under a Creative Commons BY 4.0 (CC BY 4.0) licence. Excluded material owned by third parties may include, for example, design and layout, images obtained under licence from third parties and signatures. We have made all reasonable efforts to identify and label material owned by third parties.

You may distribute, remix and build on this website's material but must attribute the AIHW as the copyright holder, in line with our attribution policy. The full terms and conditions of this licence are available at <https://creativecommons.org/licenses/by/4.0/>.

Enquiries relating to copyright should be addressed to info@aihw.gov.au.

Enquiries or comments on the METEOR metadata or download should be directed to the METEOR team at meteor@aihw.gov.au.

COVID-19 Register; Quality Statement

Identifying and definitional attributes

Metadata item type:	Data Quality Statement
Synonymous names:	COVID-19 linked data set
METEOR identifier:	788300
Registration status:	AIHW Data Quality Statements , Standard 07/03/2024

Data quality

Data quality statement summary:

Description

The COVID-19 Register brings together COVID-19 cases from states and territories and the Commonwealth Department of Health and Aged Care's National Notifiable Diseases Surveillance System (NNDSS) combined with national health administrative data sets including the:

- Australian Immunisation Register (AIR)
- Australian and New Zealand Intensive Care Society (ANZICS) Adult Patient Database (APD) and
- Australian and New Zealand (ANZ) Paediatric Intensive Care Registry (PICR)
- Medicare Benefits Schedule (MBS)
- Medicare Consumer Directory (MCD)
- National Aged Care Data Clearinghouse (NACDC)
- National Death Index (NDI)
- National Disability Insurance Scheme (NDIS)
- National Hospitals Morbidity Database (NHMD)
- National Non-Admitted Patient Emergency Department Care Database (NNAPEDCD)
- National Public Hospitals Establishment Database (NPHEd)
- Pharmaceutical Benefits Scheme (PBS, including Repatriation Schedule of Pharmaceutical Benefits (RPBS) information).

The COVID-19 Register aims to provide greater insights into the longer-term impact of COVID-19 on the health of the Australian population and the health system. Research outcomes are anticipated to inform health service planning, monitoring and evaluation purposes and policy development. The data set can be accessed by approved analysts through a secure remote environment.

Summary of Key Issues

Participation and contribution to the COVID-19 Register by jurisdictions is voluntary. The latest version (version 2.5) of this data set includes COVID-19 case notification data from New South Wales (NSW), Victoria (Vic), Queensland (Qld), South Australia (SA), Tasmania (Tas), Northern Territory (NT), and the Australian Capital Territory (ACT).

While the Australian Institute of Health and Welfare (AIHW) continues to explore avenues to secure approvals for data sharing with all jurisdictions, analysts should note limitations with the coverage of hospitals data in this version of the data set, in particular:

- Hospitalisation data from Western Australia and the Northern Territory are not included.
- Records from private hospitals are not captured in this version but will be considered in future.
- The current coverage period for most of the jurisdictions' hospitals data is from 01/07/2014–30/06/2022 with 2021–22 data for the remaining jurisdictions to be included in future updates.

The data set includes service-based data sources (such as Pharmaceutical Benefits Scheme (PBS), Medicare Benefits Schedule (MBS) and the Australian Immunisation Register (AIR)). It consists of records of services provided to people who are usual residents of Australia. It may also capture some people who live in

Australia but are not eligible for Medicare (for example, international students, visitors to Australia from countries with reciprocal healthcare agreements). In addition, everyone in Australia is eligible for a free COVID-19 vaccination and as such immunisation records will capture individuals who are not usual residents of Australia. Both under coverage and over coverage of different cohorts of interest within the Australian resident population need to be considered in the analysis and interpretation of the data set.

Each state and territory may have unique testing and reporting requirements based on jurisdictional public health orders. For jurisdictions where rapid antigen test (RAT) results are not included in case notifications linked to the COVID-19 Register, there will be an under-reporting of COVID-19 positive cases.

RATs performed in a home setting are open to user error and the obligation to record positive results rests with patients, who may not be aware of this duty, or of reporting mechanisms. This may also result in under-reporting of COVID-19 cases.

Due to under-reporting of COVID-19 RAT and/or PCR results, analysts using this data set should be aware that there might be individuals who may have never been a notified case in state and territory notifiable diseases. However, they may have died from COVID-19. These individuals will enter the cohort through the [National Deaths Index \(NDI\)](#) data source where emergency codes relating to COVID-19 were used to code cause of death.

Institutional environment:

The AIHW develops, maintains and manages the use of the COVID-19 Register and has commenced work to integrate it with the National Health Data Hub (NHDH) (currently referred to as the National Integrated Health Services Information (NIHSI)). The NHDH is an enduring linked data asset that brings together state/territory hospitals data with national health administrative data sets including data from the MBS, PBS, RPBS, Residential Aged Care services data, NDI and AIR data. Integration of the COVID-19 Register with the NHDH as one asset going forward means all data will be linked and stored once and for multiple uses, have consistent data governance arrangements, and use a consistent linkage spine. The AIHW is an independent corporate Commonwealth entity under the [Australian Institute of Health and Welfare Act 1987](#) (AIHW Act), governed by a [management board](#) and accountable to the Australian Parliament through the Health portfolio.

The AIHW is a nationally recognised information management agency. Its purpose is to create authoritative and accessible information and statistics that inform decisions and improve the health and welfare of all Australians.

Compliance with the confidentiality requirements in the AIHW Act, the Privacy Principles in the [Privacy Act 1988](#) (Cth) and AIHW's data governance arrangements ensures that the AIHW is well positioned to release information for public benefit while protecting the identity of individuals and organisations.

For further information see the AIHW website (www.aihw.gov.au/about-us), which includes details about the AIHW's governance (www.aihw.gov.au/about-us/our-governance) and our role and strategic goals (www.aihw.gov.au/about-us/what-we-do)

The establishment of the COVID-19 Register is funded by the Medical Research Future Fund (MRFF). The AIHW engages in quarterly consultations with members of the COVID-19 Data Advisory Group represented by members from state/territory health departments, the Commonwealth Department of Health and Aged Care, the National Centre for Immunisation Research and Surveillance and the Australian and New Zealand Intensive Care Society (ANZICS). The role of the COVID-19 Data Advisory Group (the Advisory Group) is to provide a wide range of expert advice to AIHW regarding the COVID-19 Register.

Timeliness:

The first version of the COVID-19 Register was completed in December 2022, with the second version (Version 2) completed in November 2023. The latest version (Version 2.5) aims to be ready for use by February 2024.

Timing of updates to the COVID-19 Register will be subject to timely provision of case notification data to the AIHW from state and territory governments, and timely access to the content data via the administrative data sets, and subject to agreement from data custodians.

The project aims to re-link information periodically to identify additional deaths, and to update data where available. ABS coded cause of death information will be

incorporated as it becomes available.

Population and coverage periods of the data sources that make up the COVID-19 Register are listed in Table 1.

Table 1. Population and coverage period for Version 2.5 of the COVID-19 Register

Data Set	State/Territory	Coverage period
Medicare Consumer Directory (MCD)	National (whole of population)	Extracted up to 31/12/2022
Pharmaceutical Benefits Scheme (PBS)	National (whole of population)	01/01/2015 – 31/03/2023
Medicare Benefits Schedule (MBS)	National (whole of population)	01/01/2015 – 31/03/2023
National Notifiable Diseases Surveillance System (NNDSS)	NSW, Vic, Qld, SA, Tas, NT, ACT (cases only)	01/01/2020 – 31/12/2022
National Death Index (NDI)	National (whole of population)	01/01/2015 – 31/03/2023
Australian Immunisation Register (AIR)	National (whole of population)	01/01/2020 – 31/03/2023
National Hospitals Morbidity Database (NHMD) - public hospitals only	NSW, Vic, Qld, SA, Tas, ACT (whole of population)	01/07/2014 – 30/06/2022 (for NSW, Vic, Qld, Tas, ACT) 01/07/2014 – 30/06/2021 (for SA only)
National Non-Admitted Patient Emergency Department Care Database (NNAPEDCD)	NSW, Vic, Qld, SA, Tas, ACT (whole of population)	01/07/2014 – 30/06/2022 (for NSW, Vic, Qld, Tas, ACT) 01/07/2014 – 30/06/2021 (for SA only)
National Public Hospitals Establishment Database (NPHEd)	Hospitals that participate in the hospitals establishment database, i.e. NSW, Vic, Qld, SA, Tas, ACT (whole of population)	01/07/2014 – 30/06/2022
Australian and New Zealand Intensive Care Society (ANZICS) Adult Patient Database (APD)	National (whole of population)	01/01/2015 – 31/12/2022

Australian and New Zealand Paediatric Intensive Care Registry (ANZPICR)	National (whole of population)	01/01/2015 – 31/12/2022
National Disability Insurance Scheme (NDIS)	National (whole of population)	01/01/2015 – 30/12/2022 (payments)
Aged Care Funding Instrument (ACFI)	National (whole of population)	01/01/2015 – 30/06/2022
National Screening and Assessment Form (NSAF)		01/01/2015 – 30/06/2022
Permanent and respite residential aged care (RAC)		01/01/2015 – 30/06/2022
Commonwealth Home Support Program (CHSP)		01/01/2015 – 30/06/2020
Home Care Packages (HCP)		01/01/2015 – 31/06/2022
Transition Care Program (TCP)		01/01/2015 – 30/06/2022
Short-Term Restorative Care (STRC)		01/01/2015 – 30/06/2022
State and territory notifiable disease databases	NSW: Notifiable Conditions Information Management System (NCIMS)	25/01/2020 – 30/09/2022
	Vic: Public Health Events Surveillance System (PHESS)/Transmission and Response Epidemiology Victoria (TREVI)	25/01/2020 – 31/12/2022
	Qld: Notifiable Conditions System (NOCS)	28/01/2020 – 20/08/2022
	SA: Notifiable Infectious Disease Database (NIDD)	30/01/2020 – 11/02/2022
	Tasmania: Tasmanian Notifiable Disease Database (TNDD)	02/03/2020 – 31/12/2022
	ACT: ACT Notifiable Diseases Register	12/03/2020 – 31/12/2022
	NT: NT Notifiable Diseases Surveillance System (NTNDS)	21/02/2020 – 31/12/2022

Accessibility:

Both government and non-government researchers are eligible to apply for access to the de-identified data in the secure remote environment. The current ethics approval only allows the data to be accessed by researchers located in Australia. If there is a need to allow access to overseas researchers in future, relevant ethics approvals will need to be obtained.

The current version of the data variable list can be viewed on the [COVID-19 Register](#) website.

Where data is not publicly available, you may request data by following the steps outlined on the [Data on request](#) page on the AIHW website in the first instance. Alternatively, the COVID-19 project team can be contacted on covid19register@aihw.gov.au. Any data request will need to be approved by the relevant data custodians. Requests that take longer than half an hour to compile will incur a charge on a cost recovery basis.

All AIHW-authored reports and publication products derived from the use of the COVID-19 Register, satisfying output requirements and approval processes will be published and accessible from the AIHW website (www.aihw.gov.au). Publications derived from external researchers will be referenced on the AIHW website once their work is published in the public domain.

Interpretability:

Information on the COVID-19 Register which can help provide insight into the data include data variable lists updated at regular intervals, fact sheets, and web reports with summary outputs derived from the linked data set. These have been made available publicly on the COVID-19 Register website in December 2022 and thereafter. Where available, metadata for each underlying data source will be published in the AIHW's online metadata registry – METEOR which can be accessed on the AIHW website at [METEOR home \(aihw.gov.au\)](#).

National notification data on COVID-19 confirmed cases is collated in the [National Notifiable Diseases Surveillance System \(NNDSS\)](#) based on notifications made to state and territory health authorities under the provisions of their relevant public health legislation. NNDSS case notifications and variables of interest are available for the 7 participating jurisdictions (NSW, Vic, Qld, SA, Tas, NT and ACT) in Version 2.5 of the COVID-19 Register. Metadata relating to state/territory and NNDSS variables used in the COVID-19 Register will be available on [METEOR \(aihw.gov.au\)](#).

The National Death Index (NDI) is maintained by AIHW and its data quality statement can be found on the AIHW website at [National Death Index \(NDI\), Data Quality Statement](#).

The data quality statements underpinning the AIHW National Mortality Database can be found in the following Australian Bureau of Statistics (ABS) publications:

- [Deaths, Australia methodology, 2022 | Australian Bureau of Statistics](#)
- [Causes of Death, Australia methodology, 2022 | Australian Bureau of Statistics](#)

Data quality statements are not specifically available, however, the metadata on [Medicare Benefits Schedule \(MBS\) data collection](#) and [Pharmaceutical Benefits Scheme \(PBS\)](#) can be found on the AIHW website.

Data quality statements for the [AIHW National Aged Care Data Clearinghouse](#) and [Aged Care Funding Instrument](#) can be found on [METEOR](#).

The data quality statements for hospitals data can be found in [About the data - Australian Institute of Health and Welfare](#).

Information on the ANZICS data sets can be found on the ANZICS website [Home - ANZICS](#).

Information on the NDIS data sets can be found on the NDIS website [Datasets | NDIS](#).

Relevance:

The COVID-19 Register was created in response to current and ongoing monitoring of the health outcomes and health system needs of people who have had a COVID-19 diagnosis.

There is potential for the COVID-19 Register to be used for research under the following approved themes:

- Epidemiological and statistical research
- Service use, medication dispensing and patient journeys
- Identifying groups or cohorts of interest
- Monitoring, evaluation and data quality improvement

The COVID-19 Register is a national, person-based linked data set, comprising at its core a register of people with a positive diagnosis of COVID-19 recorded in the first few years (depending on data supply) after the initial case was recorded in Australia.

The COVID-19 case cohort is derived from state and territory notifiable diseases databases and the NDI. It includes all people tested (via PCR and RAT) who had returned one or more positive results for COVID-19 (SARS-CoV-2 positive) at the time of data extract. That is, where the Commonwealth DISEASE_CODE = 081 (COVID-19) in state and territory notifiable diseases data, or where cause of death on the NDI identifies a COVID-19 diagnosis.

The following groups are excluded from the COVID-19 case cohort and will instead be used to define a comparison group of non-cases. That is, where the Commonwealth DISEASE_CODE is not 081 (COVID-19) in state and territory notifiable diseases data. This comparison group is derived from the Medicare Consumer Directory (MCD) and includes all people (excluding those who died before 2015) who:

- did not return a positive test result for COVID-19 at the time of data extract (these people are not included in the state and territory notifiable diseases data and therefore are excluded)
- had not been tested at the time of data extract (these people are not included in the state and territory notifiable diseases data and therefore excluded)
- are not in the notifiable diseases data cohort, who died and did not have cause of death = U07.1 "COVID-19 virus identified - used when COVID-19 is confirmed by laboratory testing" in ABS-coded cause of death information.

The reference period of each data source and list of participating jurisdictions in the Version 2.5 of the COVID-19 Register is provided in Table 1.

Data on a person's usual residence varies across data sources in the COVID-19 Register, with the minimum geography available at the postcode level or Statistical Area Level 2 (based on the ABS Australian Statistical Geography Standard (ASGS) Edition 3). SA2 level information is required to analyse data by Local Government Area (LGA), and for the derivation of socioeconomic, remoteness and other area classifications commonly used to make comparisons by region (such as SA3 and SA4).

Geographical detail at the postcode and SA2 level instead of individual addresses, are provided in the following data sources:

- SA2: MCD, NDI, NHMD, NNAPEDCD, AIR and NACDC
- Postcode: States/territories case notification, NHMD, NNAPEDCD, and NACDC

Public hospital admitted patient episode data are drawn from the National Hospital Morbidity Database (NHMD). Some admitted patients may not be enrolled in or be eligible for Medicare but will still be included in the NHMD such as international students or some overseas visitors who were admitted to public hospitals. For example, overseas visitors from New Zealand, Ireland, the United Kingdom, the Netherlands, Sweden, Finland Norway, Italy, Malta, Belgium and Slovenia may receive public hospital care because Australia has Reciprocal Health Care Agreements with these countries. Over coverage in these cases may occur due to a lack of information or when the individuals leave Australia and no longer considered as usual residents. This may mean that individuals may continue to be counted in the analysis after these individuals are no longer a resident in Australia unless, methods are applied to adjust for this over coverage. As such under coverage or over coverage of different groups within the Australian resident population need to be considered in the analysis and interpretation of data. Hospital data are reported based on state of service and not state of usual residence.

Data on admitted patient private hospitals is not available in Version 2.5 of the COVID-19 Register and as such, is unrepresented in any potential analyses and

outputs. Data custodian approval will be sought to identify organisations (hospitals) in research outputs, and no outputs identifying hospitals will be released without such approval.

Data on Indigenous status are available in the NNDSS, NHMD, NNAPEDCD, NACDC, and AIR data sources.

Accuracy:

Data included in the COVID-19 Register is sourced from AIHW data holdings for MBS, PBS, hospitals, residential aged care, and national deaths data. The data collection and cleaning processes vary across these collections, and the subsequent quality of the linked data set will be subject to the quality of the data held in these source collections. It is important for researchers to recognise that statistical outputs for analysis are generally not the primary reason for the collection of these administrative data.

Data linkage was undertaken using probabilistic linkage, involving creation of record pairs by combining records from one data set with records from another data set based on similarities in characteristics such as last name, first name(s), date of birth, sex and address of residence. Matches are evaluated based on the level of similarities between the characteristics. A higher level of similarities suggests that a given record pair is more likely to be the same person and treated as a true link.

Quality of linkage depends on the coverage and quality of identifiers available for each collection, and consistency with information held in the integrating spine (that is, MCD, NDI and AIR data were first linked to create the linkage spine).

Link accuracy for the COVID-19 Register was a high priority. The addition of AIR data to the MCD-NDI linkage spine increased the rate of successful linked individuals across all jurisdictions. Unlinked records were identified and retained in the data set. This allows analysts the opportunity to conduct sensitivity analyses using characteristics of the unlinked cohort, if required.

The AIHW conducted a program of testing and validation to ensure the integrity and quality of the data set. These checks include:

- completeness of records that is, identifying missing values or duplicates
- alignment of broad aggregates between the linked data set and the source data sets
- comparability of outputs against published sources of information.

Data that are found to be missing, duplicated or are potential outliers will be identified and provided to analysts in a user guide document to assist in their analytical processes.

Jurisdiction-specific accuracy issues include the possible inconsistencies in the way demographic variables like 'sex' is coded at data collection that is, whether this intended as sex at birth or gender and the possibility that this may not be understood by respondents.

Analysts who wish to use indigenous status data for statistical reporting purposes should note the small number of events in selected jurisdictions.

The linkage rates for the states and territories (NSW, Vic, Qld, SA, Tas, ACT and NT) data sets that were linked to the MCD-NDI-AIR spine in Version 2.5 can be found at: [Linkage results, Linkage findings - Australian Institute of Health and Welfare](#).

Coherence:

The COVID-19 Register is anticipated to be regularly updated as data become available for linkage. Differences in scope, reference periods and variables (or data sources) between each version of the COVID-19 Register will be captured in updated versions of the data quality statement. This should be considered when comparing outputs with different reference periods and versions of the COVID-19 Register.

Researchers should note that the hospitals data in the COVID-19 Register is derived from existing linkages used in the National Integrated Health Services Information Analysis Asset (NIHSIAA) project. As such, the scope of hospitals data (NHMD, NNAPEDCD and NPHEd) may be different to unlinked hospitals data.

Personal project numbers are assigned for each unique individuals in this linked data set and as such, they remain specific and relevant only to this cohort.

Demographic data, such as sex, usual residence and/or Indigenous status may be captured differently across source data collections. States/territories providing COVID-19 case notification data have different reporting conditions in capturing information on COVID-19 diagnosis (such as the capturing of RAT and/or PCR results) and COVID-19 hospitalisations or death. Work is underway to capture and understand the nuances in the reporting of these variables to be able to advise analysts when comparing data or statistics to other data collections. It is also important to note the different coverage periods between each jurisdiction in the Version 2.5 of the COVID-19 Register as this can impact comparability both within the linked data set and between other sources of data collection.

Data products

Implementation start date: 07/03/2024

Source and reference attributes

Submitting organisation: Australian Institute of Health and Welfare

Relational attributes

Related metadata references:

Supersedes [COVID-19 Register: Quality Statement](#)
[AIHW Data Quality Statements](#), Superseded 07/03/2024

See also [National Integrated Health Service Information \(NIHSI\) version 2.0](#)
[AIHW Data Quality Statements](#), Standard 21/03/2024

See also [National Integrated Health Service Information Analysis Asset \(NIHSIAA\) version 1.0](#)
[AIHW Data Quality Statements](#), Superseded 21/03/2024